Technische Universität Ilmenau
Fakultät für Mathematik
und Naturwissenschaften
Institut für Mathematik

Postfach 10 0565
98684 Ilmenau
Germany
Tel.: 03677/693621
Fax: 03677/693270
Telex: 33 84 23 tuil d.
email: W.Neundorf@mathematik.tu-ilmenau.de

Preprint No. M 5/97

Unvollständige LU-Zerlegung und approximative Inverse von Blocktridiagonalmatrizen

Werner Neundorf, Thomas Ortlepp

Juni 1997

[‡]MSC (1991): 65-04, 65F05, 65F10, 65F50, 68Q25

1 Einleitung

Die Lösung linearer Gleichungssysteme, die unter anderem bei der Diskretisierung von Differentialgleichungen entstehen, kann mittels direkter oder iterativer Verfahren erfolgen.
Treten hier große sparse Matrizen auf, so sind alle Möglichkeiten einer effizienten Implementierung auf dem Rechner einzubeziehen. Dabei sind folgende Aspekte zu berücksichtigen:

- günstige Speicherstrategien,
- Erhaltung von Strukturen in der Matrix und bei ihrer Umformung, d.h. auch Kontrolle des Auffüllungsprozesses (fill-in) bei direkten Verfahren,
- Zerlegung (Faktorisierung) der Matrix,
- näherungsweise Zerlegung der Matrix mit dem Ziel der Gewinnung von brauchbaren Vorkonditionierern für iterative Methoden,
- Bestimmung von Näherungsinversen für die Matrix bzw. Inversen von Teilstrukturen der Matrix.

Ausgefeilte Techniken nutzen die Kombination von direkten und iterativen Verfahren und damit die Vorzüge beider Varianten.
Ziel dieser Arbeit soll es sein, die unvollständige Zerlegung zusammen mir der Idee der approximativen Inversen für einen speziellen Typ von tridiagonalen Matrizen zu untersuchen und aufzuzeigen, wie weit verschiedene Ansätze dabei genutzt werden können. Weiterführende Betrachtungen und Literaturhinweise findet man in [1].

2 Ein Beispiel

Folgendes Beispiel wird auch bei den verschiedenen Tests angewandt.
Bei der üblichen Diskretisierung von 2D-Problemen partieller Differentialgleichungen entstehen in der Regel Matrizen mit einer Blocktridiagonalstuktur. So ergibt sich für das ebenene stationäre Wärmeleitproblem mit entsprechenden Randbedingungen in einem rechteckigen Gebiet bei äquidistanter Vernetzung mit 3 x 4 inneren Stützstellen die Matrix

$$M = \begin{pmatrix} \begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix} & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & & \\ \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix} & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \\ & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix} & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} \\ & & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix} \end{pmatrix}.$$

Wenn die Anzahl der Stützstellen praktische Größenordnungen annimmt, so werden die entstehenden Gleichungssysteme oft mittels iterativer Verfahren gelöst. Der Aufwand für eine direkte Lösung entspricht dem einer exakten LU-Zerlegung der Matrix.

Was die Iteration betrifft, will man eine schnelle Konvergenz erreichen, so daß eine problemspezifische Vorkonditionierung eine wichtige Rolle spielt. Dabei kann man sich mit einer näherungsweisen LU-Zerlegung zufriedengeben, die natürlich mit weniger Aufwand zu beschaffen ist und zur Beschreibung des Vorkonditionierers genommen wird.

3 Die LU-Zerlegung einer Tridiagonalmatrix

Betrachten wir das tridiagonale Gleichungssystem $Ax = f,\ A \in \mathbb{R}^{(n,n)},\ x, f \in \mathbb{R}^n$, wobei die Matrix A die folgende Form hat

$$A = \begin{pmatrix} a_1 & b_1 & & & & \\ c_2 & a_2 & b_2 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & b_{n-1} \\ & & & & c_n & a_n \end{pmatrix} = tridiag(c_i, a_i, b_i).$$

Def.: Ein Tridiagonalmatrix $A = tridiag(c_i, a_i, b_i)$ der Dimension n heißt *irreduzibel diagonaldominant*, wenn die Bedingungen (i) bis (iii) erfüllt sind.

$$\begin{array}{lll} (i) & |a_1| > |b_1| > 0, & \\ (ii) & |a_i| \geq |c_i| + |b_i|, \quad c_i \neq 0, b_i \neq 0 & \text{für} \quad 2 \leq i \leq n-1, \\ (iii) & |a_n| \geq |c_n| > 0. & \end{array}$$

Die Definition ist in [2] angegeben und stellt einen speziellen Fall ihrer allgemeinen Version dar. An gleicher Stelle wird gezeigt, daß die obigen Bedingungen hinreichend sind für die Existenz und Eindeutigkeit der Lösung des Gleichungssystems. Die Lösbarkeit ist auch noch unter leicht abgeschwächten Voraussetzungen gegeben. Das klassische Eliminationsverfahren von Gauß führt zu dem folgenden speziellen Algorithmus ($\mathcal{A}$) der Matrixzerlegung.

(1) $\alpha_1 := a_1$

(2) $\gamma_1 := \alpha_1^{-1} * b_1$

(3) für $i = 2, 3, ..., n-1$
$\alpha_i := a_i - c_i * \gamma_{i-1}$
$\gamma_i := \alpha_i^{-1} * b_i$

(4) $\alpha_n := a_n - c_n * \gamma_{n-1}$

Die Matrizen der Dreieckszerlegung $A = LU$ sind somit

$$L = tridiag(c_i, \alpha_i, 0), \quad U = tridiag(0, 1, \gamma_i).$$

Man bemerke, daß die exakte Dreieckszerlegung (Faktorisierung) natürlich auch die Übereinstimmung der Zeilensummen von A und LU - analog Spaltensummen - gewährleistet. Sie wird deswegen auch als *tridiagonale Identifikation* bezeichnet:

Zeilen: $c_i + a_i + b_i = c_i + (c_i\gamma_{i-1} + \alpha_i) + \alpha_i\gamma_i,$
Spalten: $b_{i-1} + a_i + c_{i+1} = \alpha_{i-1}\gamma_{i-1} + (c_i\gamma_{i-1} + \alpha_i) + c_{i+1}.$

Anstelle von 2 Vektoren α, γ kann man die Zerlegung auch auf die Berechnung eines Vektors zurückführen. Es gilt

$$A = LU = LD^{-1}U' = tridiag(c_i, m_i, 0)\ diag(m_i^{-1})\ tridiag(0, m_i, b_i)$$

mit $D = diag(m_i),\ m_1 = a_1,\ m_i = a_i - c_i b_{i-1}/m_{i-1},\ i = 2, ..., n.$
In Bezug auf das obige Beispiel mit blocktridiagonaler Struktur erhält man sofort eine Zerlegungsmethode, wenn man im Algorithmus $(\mathcal{A})$ anstelle der Matrixelemente die Blockmatrizen notiert. Dabei ist natürlich die Division durch die Inversion der Teilmatrizen zu ersetzen.

4 Die LU-Zerlegung einer Blocktridiagonalmatrix

Die Beispielmatrix habe die Dimension (nm, nm), wobei die einzelnen Teilblöcke der Dimension (n, n) von tri- bzw. diagonaler Struktur sind.
Wir führen folgende Bezeichnung ein: $I = I(n, n)$ Einheitsmatrix der Dimension n,

$$A = \begin{pmatrix} D_1 & U_1 & & & \\ L_2 & D_2 & U_2 & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & U_{m-1} \\ & & & L_m & D_m \end{pmatrix} = \begin{pmatrix} T_1 & & & & \\ L_2 & T_2 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & L_m & T_m \end{pmatrix} \begin{pmatrix} I & W_1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & W_{m-1} \\ & & & & I \end{pmatrix}.$$

Der Algorithmus nimmt nun die Form $(\mathcal{B})$ an.

(1) $T_1 := D_1$

(2) $H := inverse(T_1)$
$W_1 := H * U_1$

(3) für $j = 2, 3, ..., m-1$
$T_j := D_j - L_j * W_{j-1}$
$H := inverse(T_j)$
$W_j := H * U_j$

(4) $T_m := D_m - L_m * W_{m-1}$

Das Hauptproblem im Algorithmus $(\mathcal{B})$ liegt in der Art und Weise der Bestimmung der inversen Matrix und ihrer jeweiligen Verarbeitung im nächsten Schritt.

Hierbei sind zwei Aspekte von Bedeutung. Zum einen stellt die Berechnung von T_j^{-1} einen erheblichen Zeitaufwand dar, man denke sich z.B. $n = m = 1000$, was auf 1000 Inversionen von Matrizen der Größe $(1000, 1000)$ führt. Andererseits ist die Inverse einer Tridiagonalmatrix eine voll besetzte Matrix. Damit sind es auch $W_1, ..., W_{m-1}, T_2, ..., T_m$. Somit entsteht ein erheblicher Speicherbedarf.
Da aber die Zerlegung zur Vorkonditionierung für ein Iterationsverfahren verwendet werden soll, wäre eine approximative LU-Zerlegung auch schon von Nutzen.

5 Approximative Inverse einer Tridiagonalmatrix

Sei $A = tridiag(c_i, a_i, b_i)$ eine irreduzibel diagonaldominante Matrix.
Wir versuchen nun, eine Matrix $Z \approx A^{-1}$, die eine bestimmte Struktur aufweisen soll, so zu ermitteln, daß

$$||I - ZA|| \to min$$

wird.
Im allgemeinen kann man Z nach mehreren Gesichtspunkten auswählen.

$$Z = \begin{cases} \text{linke oder rechte approximative Inverse,} \\ \text{Besetzungsstruktur (diagonal, tridiagonal,...),} \\ \text{Kriterium der Approximation (Frobenius-Norm, Spektraläquivalenz,...).} \end{cases}$$

Obige Minimierungsaufgabe verwendet die approximative Inverse von links. Als Norm sei wegen daraus sich ergebender Differenzierbarkeit die Frobenius-Norm

$$||B||_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}^2$$

gewählt.

5.1 Approximation der Inversen durch eine Diagonalmatrix

Der erste Ansatz für die approximative Inverse sei $Z = diag(q_i)$.
Dann gilt

$$F(q) = ||I - ZA||_F^2 = \sum_{i=1}^{n} [q_i^2 c_i^2 + (1 - q_i a_i)^2 + q_i^2 b_i^2], \quad q = (q_1, ..., q_n), \; c_1 = b_n = 0.$$

Mittels der notwendigen Bedingung für das Minimum folgt

$$\begin{aligned} 0 &= \frac{\partial F}{\partial q_i} = 2q_i c_i^2 + 2(1 - q_i a_i)(-a_i) + 2q_i b_i^2, \\ 0 &= q_i(c_i^2 + a_i^2 + b_i^2) - a_i, \\ q_i &= \frac{a_i}{c_i^2 + a_i^2 + b_i^2} \quad \text{für} \quad i = 1, ..., n. \end{aligned}$$

Die so definierte Matrix Z ist eine im obigen Sinne gute Näherung der Inversen von A, insbesondere bei starker Diagonaldominanz von A.
Für den Fall, daß A die Form des Diagonalblocks D_j der Beispielmatrix hat, läßt sich der Fehler der Approximation abschätzen.

$$q_1 = q_n = \frac{4}{17}, \quad q_i = \frac{a_i}{c_i^2 + a_i^2 + b_i^2} = \frac{4}{18}, \quad i = 2, 3, ..., n-1$$

$$||I - ZA||_F^2 = \left\| I - \begin{pmatrix} \frac{4}{17} & & & & \\ & \frac{4}{18} & & & \\ & & . & & \\ & & & \frac{4}{18} & \\ & & & & \frac{4}{17} \end{pmatrix} \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & . & . & . & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{pmatrix} \right\|_F^2$$

$$= \left\| \begin{pmatrix} \frac{1}{17} & \frac{4}{17} & & & \\ \frac{2}{9} & \frac{1}{9} & \frac{2}{9} & & \\ & . & . & . & \\ & & \frac{2}{9} & \frac{1}{9} & \frac{2}{9} \\ & & & \frac{4}{17} & \frac{1}{17} \end{pmatrix} \right\|_F^2$$

$$= (n-2)\left(\frac{1}{9}\right)^2 + 2\left(\frac{1}{17}\right)^2 + 2(n-2)\left(\frac{2}{9}\right)^2 + 2\left(\frac{4}{17}\right)^2$$

$$||I - ZA||_F^2 = \frac{n}{9} - \frac{16}{153}$$

$$||A||_F^2 = 4^2 n + 2(n-1)(-1)^2 = 18n - 2$$

$$\lim_{n\to\infty} \frac{||I - ZA||_F}{||A||_F} = \lim_{n\to\infty} \sqrt{\frac{\frac{n}{9}}{18n}} = \frac{1}{\sqrt{162}} \approx 0.078$$

Der relative Fehler bleibt also beschränkt. Seine Größe hängt direkt von der Diagonaldominanz der Matrix A ab. Je ausgeprägter diese ist, desto kleiner ist der relative Fehler. Der Fehler wird Null, wenn A selbst Diagonalmatrix ist.
Nimmt man z.B. die Matrix $A' = tridiag(-b, a, -b)$, $a \gg b > 0$, so kann man in analoger Weise nachrechnen, daß

$$\lim_{n\to\infty} \frac{||I - ZA'||_F}{||A'||_F} = \frac{\sqrt{2}\, b}{a^2 + 2b^2}$$

Zum Vergleich sei noch eine andere diagonale approximative Inverse angegeben, und zwar ihre einfache Version $Z = diag(a_i^{-1}) = diag(\frac{1}{4})$, die die Matrix $I - ZA = tridiag(\frac{1}{4}, 0, \frac{1}{4})$ und den Fehler $||I - ZA||_F^2 = \frac{n}{8} - \frac{1}{8}$ liefert. Letzterer ist größer als der optimale Fehler $\frac{n}{9} - \frac{16}{153}$ bei $n > 1$.

Verwendet man die Matrix $Z = diag(q_i)$ als Vorkonditionierer im Iterationsverfahren (Fixpunktiteration)

$$x^{(k+1)} = x^{(k)} - Z(Ax^{(k)} - f) = (I - ZA)x^{(k)} + Zf,$$

so nützt einem bei der Untersuchung seiner Konvergenz die Normabschätzung $||I - ZA||_F \approx \frac{\sqrt{n}}{3}$ natürlich wenig. Auch die Äquivalenzaussagen zur Frobeniusnorm von $B = I - ZA$ mit Zeilen- bzw. Spaltensummennorm gemäß

$$\frac{1}{3} \approx \frac{1}{\sqrt{n}}||B||_F \leq ||B||_{\infty,1} \leq \sqrt{n}||B||_F \approx \frac{n}{3}$$

sind zu grob, um z.B. auf $||B||_\infty < 1$ zu schließen. Nach obigen Berechnungen folgt aber $||B||_\infty = \frac{5}{9} < 1$.

5.2 Approximation der Inversen durch eine Tridiagonalmatrix

Der nächste Ansatz für die approximative Inverse sei $Z = tridiag(x_i, z_i, y_i)$.
Die Herleitung seiner Komponenten erfolgt analog zum obigen Fall auf der Basis der notwendigen Bedingungen zur Minimierung des Funktionals

$$F(x_2, ..., x_n, z_1, z_2, ..., z_n, y_1, ..., y_{n-1}) = ||I - ZA||_F^2 =$$

$$= \sum_{i=1}^{n} [x_i^2 c_{i-1}^2 + (x_i a_{i-1} + z_i c_i)^2 + (1 - x_i b_{i-1} - z_i a_i - y_i c_{i+1})^2 + (z_i b_i + y_i a_{i+1})^2 + y_i^2 b_{i+1}^2]$$

$(x_1 = y_n = 0)$ gemäß

$$0 = \frac{\partial F}{\partial x_i}, \quad 0 = \frac{\partial F}{\partial z_i}, \quad 0 = \frac{\partial F}{\partial y_i}.$$

Für jedes $i = 1, 2, ..., n$ entsteht eine 3x3 Gleichungssystem der Form

$$\begin{pmatrix} c_{i-1}^2 + a_{i-1}^2 + b_{i-1}^2 & c_i a_{i-1} + a_i b_{i-1} & c_{i+1} b_{i-1} \\ c_i a_{i-1} + a_i b_{i-1} & c_i^2 + a_i^2 + b_i^2 & c_{i+1} a_i + a_{i+1} b_i \\ c_{i+1} b_{i-1} & c_{i+1} a_i + a_{i+1} b_i & c_{i+1}^2 + a_{i+1}^2 + b_{i+1}^2 \end{pmatrix} \begin{pmatrix} x_i \\ z_i \\ y_i \end{pmatrix} = \begin{pmatrix} b_{i-1} \\ a_i \\ c_{i+1} \end{pmatrix},$$

wobei $c_0 = c_1 = a_0 = b_0 = b_n = b_{n+1} = a_{n+1} = c_{n+1} = 0$ sind.
Man bemerke, daß seine Koeffizientenmatrix A_i symmetrisch und $diag(A_i) > 0$ sind. Ist A irreduzibel diagonaldominant, so stellt die Lösung des 3x3-Systems kein Problem dar.
Auch hier kann der Fehler für die Beispielmatrix $D_j = tridiag(-1, 4, -1)$ wieder bestimmt werden. Die lokalen Gleichungssyteme für $i = 3, 4, ..., n-2$ lauten

$$\begin{pmatrix} 18 & -8 & 1 \\ -8 & 18 & -8 \\ 1 & -8 & 18 \end{pmatrix} \begin{pmatrix} x \\ z \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ -1 \end{pmatrix},$$

und haben die Lösung

$$\begin{pmatrix} x \\ z \\ y \end{pmatrix} = \frac{1}{107} \begin{pmatrix} 7 \\ 30 \\ 7 \end{pmatrix}.$$

Zur Vollständigkeit noch die beiden ersten und letzten Systeme (für $i = 1, n-1$ haben diese die reduzierte Dimension 2) mit ihren Lösungen:

$i = 1$:

$$\begin{pmatrix} 17 & -8 \\ -8 & 18 \end{pmatrix} \begin{pmatrix} z_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} z_1 \\ y_1 \end{pmatrix} = \frac{1}{242} \begin{pmatrix} 64 \\ 15 \end{pmatrix},$$

$i = 2$:

$$\begin{pmatrix} 17 & -8 & 1 \\ -8 & 18 & -8 \\ 1 & -8 & 18 \end{pmatrix} \begin{pmatrix} x_2 \\ z_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} x_2 \\ z_2 \\ y_2 \end{pmatrix} = \frac{1}{1689} \begin{pmatrix} 119 \\ 478 \\ 112 \end{pmatrix},$$

$i = n-1$:

$$\begin{pmatrix} x_{n-1} \\ z_{n-1} \\ y_{n-1} \end{pmatrix} = \frac{1}{1689} \begin{pmatrix} 112 \\ 478 \\ 119 \end{pmatrix},$$

$i = n$:

$$\begin{pmatrix} x_n \\ z_n \end{pmatrix} = \frac{1}{242} \begin{pmatrix} 15 \\ 64 \end{pmatrix}.$$

Damit ist

$$Z = \begin{pmatrix} \frac{64}{242} & \frac{15}{242} & & & & & \\ \frac{119}{1689} & \frac{478}{1689} & \frac{112}{1689} & & & & \\ & \frac{7}{107} & \frac{30}{107} & \frac{7}{107} & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \frac{7}{107} & \frac{30}{107} & \frac{7}{107} & \\ & & & & \frac{112}{1689} & \frac{478}{1689} & \frac{119}{1689} \\ & & & & & \frac{15}{242} & \frac{64}{242} \end{pmatrix}.$$

Die Matrix ZA ist pentagonal und hat in den Zeilen $3, 4, ..., n-2$ die 5 Nichtnullelemente $[-\frac{7}{107}, -\frac{2}{107}, \frac{106}{107}, -\frac{2}{107}, -\frac{7}{107}]$.

Entscheidend für die Grenzwertbetrachtung in $||I - ZA||_F$ sind genau diese mittleren Zeilen von $B = I - ZA$, also $[\frac{7}{107}, \frac{2}{107}, \frac{1}{107}, \frac{2}{107}, \frac{7}{107}]$, die den Anteil $\sum_{i=3}^{n-2} b_{ij}^2 = \frac{n-4}{107}$ liefern.

Somit ergibt sich für die den relativen Fehler

$$\lim_{n\to\infty} \frac{||I - ZA||_F}{||A||_F} = \lim_{n\to\infty} \sqrt{\frac{\frac{n-4}{107}}{18n}} = \frac{1}{\sqrt{1926}} \approx 0.023$$

Der Fehler bei diesem Ansatz ist also weniger als ein Drittel des Fehlers beim "diagonalen" Verfahren. Jedoch muß man sich ihn mit dem etwa vierfachen Aufwand erkaufen.

5.3 Weitere Möglichkeiten und Analyse

Nun kann man für die Matrix Z sicherlich immer mehr Freiheitsgrade vorgeben. Hier sei noch den Ansatz für eine Pentagonalmatrix angeführt. Dieser Fall wird jedoch in den späteren Verfahren (Kap.6) nicht berücksichtigt.

Die Pentagonalform ist

$$\begin{pmatrix} z_1 & y_1 & w_1 & & & & & \\ x_2 & z_2 & y_2 & w_2 & & & & \\ u_3 & x_3 & . & . & . & & & \\ & u_4 & . & . & . & . & & \\ & & . & . & . & . & . & \\ & & & . & . & . & . & w_{n-2} \\ & & & & . & . & . & y_{n-1} \\ & & & & & u_n & x_n & z_n \end{pmatrix}.$$

Das entstehende 5x5-System hat die Koeffizientenmatrix

$$\begin{pmatrix} c_{i-2}^2 + a_{i-2}^2 + b_{i-2}^2 & a_{i-2}c_{i-1} + b_{i-2}a_{i-1} & b_{i-2}c_i & 0 & 0 \\ a_{i-2}c_{i-1} + b_{i-2}a_{i-1} & c_{i-1}^2 + a_{i-1}^2 + b_{i-1}^2 & c_i a_{i-1} + a_i b_{i-1} & b_{i-1}c_{i+1} & 0 \\ b_{i-2}c_i & c_i a_{i-1} + a_i b_{i-1} & c_i^2 + a_i^2 + b_i^2 & c_{i+1}a_i + a_{i+1}b_i & b_i c_{i+2} \\ 0 & b_{i-1}c_{i+1} & c_{i+1}a_i + a_{i+1}b_i & c_{i+1}^2 + a_{i+1}^2 + b_{i+1}^2 & c_{i+2}a_{i+1} + a_{i+2}b_{i+1} \\ 0 & 0 & b_i c_{i+2} & c_{i+2}a_{i+1} + a_{i+2}b_{i+1} & c_{i+2}^2 + a_{i+2}^2 + b_{i+2}^2 \end{pmatrix}.$$

Der Vektor der rechten Seite lautet $(0, b_{i-1}, a_i, c_{i+1}, 0)^T$.

Dazu sind wiederum die Besonderheiten bei den ersten und letzten Gleichungssystemen zu berücksichtigen.
Der Aufwand bleibt in derselben Größenordnung, wie in den anderen beiden Ansätzen. Er erhöht sich lediglich um eine Konstante.

In der Abb.1 zeigen die 3 Kurven das Verhalten des relativen Fehlers für den Diagonalblock $A = D_j$ der Beispielmatrix M über der Dimension $n = 1..95$. Von oben nach unten sind hier die Ansätze für Z diagonal, tridiagonal und pentagonal dargestellt. Die Fehlerbetrachtung hat gezeigt, daß der relative Fehler sich mit wachsenden n einer oberen Schranke nähert.

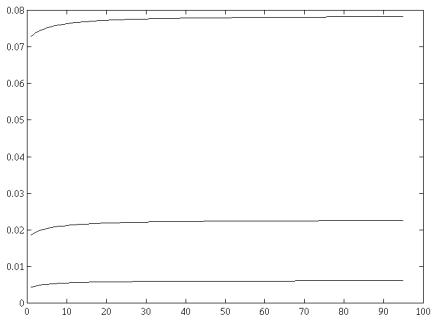

Abb.1

In der Praxis kann die Matrix eine andere, aber ähnliche Struktur aufweisen. Zu diesem Zweck sind die drei Verfahren auf diagonaldominante bzw. dazu ähnliche Matrizen mit zufälligen Störungen angewendet worden.
In Abb.2 hat A die Form $tridiag(Rand(1)+1, -4-Rand(4), Rand(1)+1)$, wobei $Rand(x)$ ein Generator für gleichverteilte Zufallszahlen in $[0,x)$ ist.
In Abb.3 sind die Störungen noch größer. Die Matrix hat die Form
$A = tridiag(Rand(2)+1, -4-Rand(4), Rand(2)+1)$.
Die Rechnungen und graphischen Darstellungen erfolgten mittels *MATLAB*.

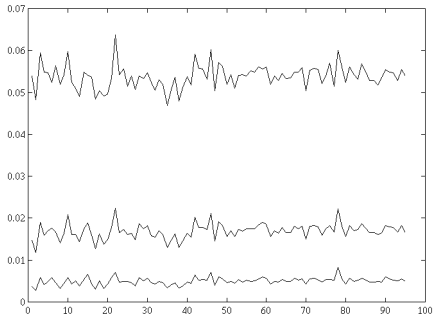

Abb.2

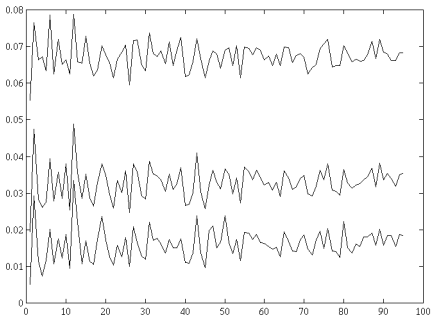

Abb.3

6 Die unvollständige LU-Zerlegung

Kehren wir zum Algorithmus ($\mathcal{B}$) aus Kap.4 zurück.
Hier besteht das Problem in der Invertierung der Tridiagonalmatrizen. Neben der exakten Inversen verwende man zunächst die zwei ersten in Kap.5 vorgestellten Appoximationen mit dem Ziel, daß sowohl W_j als auch T_j eine einfache Struktur erhalten. Hierbei ist ein Fehleranwachsen durch die $m-1$ Schritte im Verfahren zu erwarten. Des weiteren werden zum Vergleich noch 2 andere Varianten gezeigt.

(a) Exakte Inversion und vollständige Speicherung der Matrizen.

(b) Inverse approximiert durch eine Diagonalmatrix (5.1) und Speicherung der T_j als Tridiagonal-, der W_j als Diagonalmatrizen.

(c) Inverse approximiert durch eine Tridiagonalmatrix (5.2) und Speicherung aller Ergebnisse als Tridiagonalmatrizen.

(d) Exakte Inversion, aber Speicherung der T_j als Tridiagonal-, der W_j als Diagonalmatrizen. Das heißt, die Berechnung verläuft wie in (a), aber von den Zwischenergebnissen (voll besetzte Matrizen) werden nur ausgewählte Diagonalen gespeichert und später zur Kontrollrechnung herangezogen.

(e) Exakte Inversion, aber Speicherung aller Ergebnisse als Tridiagonalmatrizen.

Die Varianten (b-e) liefern genäherte Blockdreiecksmatrizen L, U.

In den Verfahren (b) und (d) wird davon ausgegangen, daß die Matrizen L_j und U_j lediglich Diagonalmatrizen sind.
Mit den verschiedenen Modifikationen des Zerlegungsverfahrens wurden Tests mit der quadratischen Beispielmatrix $M(n^2, n^2)$ durchgeführt. Dazu wurde ein PC mit Prozessor iPentium 133MHz unter Linux mit dem GNU C++ Compiler 2.6.3, Gleitkommaformat *double* (8 Byte), verwendet.

Die Zahlenwerte in der Tabelle 1 sind von oben nach unten gelesen:
- *stor* Speicherbedarf für Ergebnisdaten T_j, W_j in KBytes,
- *err* relativer Fehler gemäß

$$err = \frac{||A - LU||_F}{||A||_F},$$

- *time* Rechenzeit in Sekunden.

Verfahren	(a)	(b)	(c)	(d)	(e)
Speicherplätze für T_j, W_j, Anzahl der GK-Zahlen	$2n^2m$	$4nm$	$6nm$	$2nm$	$6nm$
Gleitkommaoperationen	$\frac{3}{5}n^3m$	$15nm$	$57nm$	$\frac{3}{5}n^3m$	$\frac{3}{5}n^3m$
$stor$:	137	13	20	7	21
$n = 20$ err :	10^{-16}	0.0805	0.0307	0.1484	0.0499
$time$:	-	-	-	-	-
	2108	86	124	76	172
$n = 50$	10^{-15}	0.1273	0.0495	0.2443	0.0865
	14.85	0.15	0.57	14.10	14.30
	16402	365	502	278	668
$n = 100$	10^{-15}	0.1802	0.0704	0.3506	0.12437
	224.20	0.56	2.13	219.10	220.00
	-	1360	2008	-	-
$n = 200$	-	0.2549	0.0999	-	-
	-	1.90	7.4	-	-
	-	5325	7866	-	-
$n = 400$	-	0.3605	0.1416	-	-
	-	7.64	28.10	-	-
	-	8330	12452	-	-
$n = 500$	-	0.4031	0.1583	-	-
	-	11.7	43.5	-	-

Tab.1. Computertest zur LU-Zerlegung von M $(m = n)$.

In den nicht gerechneten Fällen kann man die Größenordnung der Ergebnisse durch Extrapolation gewinnen.
Würde man in den Verfahren (d),(e) die nächste Berechnung der Inversen jeweils nur mit den gespeicherten Daten durchführen, dann wäre ein noch größerer Fehler zu erwarten.
Im Vergleich der Näherungsverfahren schneidet (c) am besten ab. Der Rechenaufwand ist vertretbar, und es erzielt die besten Ergebnisse. Die günstige Rechenzeit resultiert nicht nur aus der geringen Anzahl an Operationen, sondern auch aus der kompakten Speicherung der Daten im Arbeitspeicher. Jede Matrix ist als eindimensionales Feld in der Reihenfolge Super-, Haupt- und Subdiagonale abgelegt. Damit entstehen Blöcke von Vektoren für die Anteile U, D, L sowie die Ergebnisse T, W.
Für die vorliegenden Beispielrechnungen hat der zur Verfügung stehende RAM-Speicher gereicht. Wäre das unter Linux nicht der Fall, würde ein RAM-Swapping automatisch durchgeführt, das sich in einer Erhöhung der Rechenzeit auswirkt. Aufgrund der sequentiellen Struktur des Algorithmus könnte man in jedem Schritt die Teilergebnisse (hier T_j, W_j) auslagern. Das wäre dann sicherlich zeitgünstiger zu lösen als das Swapping.

6.1 Geschachtelte unvollständige LU-Zerlegung

Bisher basierte die Berechnung einer näherungsweisen LU-Zerlegung von
$A = blocktridiag(L_j, D_j, U_j) \equiv (L_j, D_j, U_j),\ D_j = tridiag(c_{ij}, a_{ij}, b_{ij}) \equiv (c_{ij}, a_{ij}, b_{ij})$,
L_j, U_j Diagonalmatrizen, auf der lokalen Approximation der Inversen einer Tridiagonalmatrix mittels diagonaler oder tridiagonaler Struktur. Dabei wurde der "tridiagonale Rahmen"

$$T_j := D_j - L_j * approx(T_j^{-1}) * U_{j-1}$$

unter den Voraussetzungen über die Gestalt der Blockmatrizen D_j, L_j, U_j insgesamt nicht gesprengt.

In [1] ist ein weiterer Zugang beschrieben worden, der eine Blockzerlegung

$$Q = blocktridiag(L_j, T_j, 0) * blocktridiag(0, I_j, W_j) \equiv (L_j, T_j, 0) * (0, I_j, W_j)$$

verwendet und dabei die Diagonalblöcke T_j als faktorisierte Tridiagonalmatrizen der Form

$$T_j = tridiag(c_{ij}, m_{ij}, 0) * tridiag(0, 1, q_{ij})$$

sucht. Zusätzlich hat man in Q bzgl. der Stuktur und Elemente der Blöcke W_j noch gewisse Freiheiten.
Nach Umformung erhält man

$$\begin{aligned} Q &= (L_j, L_j W_{j-1} + T_j, T_j W_j), \\ T_j &= (c_{ij}, c_{ij} q_{i-1,j} + m_{ij}, m_{ij} q_{ij}) = (c_{ij}, \tilde{a}_{ij}, \tilde{q}_{ij}), \\ & \quad \tilde{a}_{ij} = c_{ij} q_{i-1,j} + m_{ij}, \quad \tilde{q}_{i-1,j} = m_{i-1,j} q_{i-1,j}. \end{aligned}$$

Zu bestimmen sind also

- die Größen m_{ij}, q_{ij},
- die Struktur von W_j und seine Elemente.

Die Kriterien für eine gute globale Approximation $Q \approx A$ sind

- Spaltensummenübereinstimmung von A und Q
- Zeilensummenübereinstimmung von A und Q

jeweils mit Diagonalstruktur von W_j.

Der Algorithmus wird als geschachtelte unvollständige LU-Zerlegung (nested incomplete LU-factorization, NILU) bezeichnet.

6.1.1 NILU mit Spaltensummenkriterium

Spaltensummenübereinstimmung von A und Q heißt

$$\sum_{i=1}^{N} A_{ij} = \sum_{i=1}^{N} Q_{ij} \quad \forall j.$$

Sei $E_j^T = (1, 1, ..., 1)$ der Einsvektor mit der Länge entsprechend der Dimension der Blöcke U_{j-1}, D_j, L_{j+1}. Damit berechnet man die Spaltensummen von

A : $E_{j-1}^T U_{j-1} + E_j^T D_j + E_{j+1}^T L_{j+1}$,
Q : $E_{j-1}^T T_{j-1} W_{j-1} + E_j^T (L_j W_{j-1} + T_j) + E_{j+1}^T L_{j+1}$.

Übereinstimmung ist zu erreichen, falls

$$\begin{aligned} E_{j-1}^T U_{j-1} &= E_{j-1}^T T_{j-1} W_{j-1}, \\ E_j^T D_j &= E_j^T (L_j W_{j-1} + T_j), \quad \text{bzw.} \\ (*) \qquad E_j^T T_j &= E_j^T (D_j - L_j W_{j-1}), \\ E_j^T T_j W_j &= E_j^T U_j. \end{aligned}$$

Damit sind jedoch die Matrizen T_j, W_j noch nicht vollständig beschrieben.

Auswahl der Struktur von W_j

(1) $W_j = diag(w_{ij})$

Aus den Bedingungen $(*)$ können folgende Beziehungen abgeleitet werden.

$$\begin{aligned} D_j - L_j W_{j-1} &= (c_{ij}, a_{ij} - l_{ij} w_{i,j-1}, b_{ij}), \\ E_j^T (D_j - L_j W_{j-1}) &= [b_{i-1,j} + a_{ij} - l_{ij} w_{i,j-1} + c_{i+1,j}], \\ &\quad n \text{ Spaltensummen bilden Zeilenvektor}, \\ E_j^T T_j &= [\tilde{q}_{i-1,j} + \tilde{a}_{ij} + c_{i+1,j}], \\ \\ \tilde{q}_{i-1,j} + \tilde{a}_{ij} &= b_{i-1,j} + a_{ij} - l_{ij} w_{i,j-1}, \\ m_{i-1,j} q_{i-1,j} + c_{ij} q_{i-1,j} + m_{ij} &= b_{i-1,j} + a_{ij} - l_{ij} w_{i,j-1}. \end{aligned}$$

Da daraus noch nicht eindeutig die Größen m, q, w zu bestimmen sind, wird folgende Identifikation durchgeführt.

$$\begin{aligned} c_{ij}q_{i-1,j} + m_{ij} &= a_{ij} - l_{ij}w_{i,j-1}, \\ m_{i-1,j}q_{i-1,j} &= b_{i-1,j}. \end{aligned}$$

Aus der anderen Gleichung $E_j^T T_j W_j = E_j^T U_j$ erhalten wir

$$\begin{aligned} (\tilde{q}_{i-1,j} + \tilde{a}_{ij} + c_{i+1,j})w_{ij} &= u_{ij}, \text{ d.h.} \\ (m_{i-1,j}q_{i-1,j} + m_{ij} + c_{ij}q_{i-1,j} + c_{i+1,j})w_{ij} &= u_{ij}. \end{aligned}$$

Zusammenfassung der 3 Bedingungen:

$j = 1, 2, ..., m$

$$\left\{ \begin{array}{rcl} m_{ij} &=& a_{ij} - c_{ij}q_{i-1,j} - l_{ij}w_{i,j-1} \\ m_{ij}q_{ij} &=& b_{ij} \end{array} \right\} \quad i = 1, 2, ..., n-1, (n),$$

$$(m_{i-1,j}q_{i-1,j} + m_{ij} + c_{ij}q_{i-1,j} + c_{i+1,j})w_{ij} = u_{ij}, \quad i = 1, 2, ..., n.$$

Zuerst werden also m_{ij}, q_{ij} für alle i, dann w_{ij} ermittelt. $w_{i,0}, q_{0,j}, m_{0,j}, c_{n+1,j}$ verschwinden.
Der Aufwand des Algorithmus beschränkt sich für jedes j auf skalare Rechnungen. Damit fällt die Approximation $A \approx Q$ relativ grob aus. Vergleicht man z.B. ihre jeweils letzten Anteile U_j und $T_j W_j$, so wäre bei $W_j = T_j^{-1} U_j$ der diagonale Ansatz für W_j garnicht realistisch.

Beispielmatrix: $m = 3,\ n = 4,\ M \approx Q = (L_j, L_j W_{j-1} + T_j, T_j W_j)$

$$Q = \begin{pmatrix} \begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix} & & & \\ \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} \frac{11}{3} & -1 & \\ -1 & \frac{7}{2} & -1 \\ & -1 & \frac{11}{3} \end{pmatrix} & & \\ & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} \frac{29}{8} & -1 & \\ -1 & \frac{10}{3} & -1 \\ & -1 & \frac{29}{8} \end{pmatrix} & \\ & & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} \frac{76}{21} & -1 & \\ -1 & \frac{13}{4} & -1 \\ & -1 & \frac{78}{21} \end{pmatrix} \end{pmatrix} *$$

$$\begin{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} & \begin{pmatrix} -\frac{1}{3} & & \\ & -\frac{1}{2} & \\ & & -\frac{1}{3} \end{pmatrix} & & \\ & \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} & \begin{pmatrix} -\frac{3}{8} & & \\ & -\frac{2}{3} & \\ & & -\frac{3}{8} \end{pmatrix} & \\ & & \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} & \begin{pmatrix} -\frac{8}{21} & & \\ & -\frac{3}{4} & \\ & & -\frac{8}{21} \end{pmatrix} \\ & & & \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \end{pmatrix}.$$

Alle Diagonalblöcke $T_1,\ L_jW_{j-1}+T_j,\ j=2,3,4$, von Q haben die Form D_1. Somit ist

$$Q=\begin{pmatrix}
\begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix} & \begin{pmatrix} -\frac{4}{3} & \frac{1}{2} & \\ \frac{1}{3} & -2 & \frac{1}{3} \\ & \frac{1}{2} & -\frac{4}{3} \end{pmatrix} & & \\
\begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix} & \begin{pmatrix} -\frac{11}{8} & \frac{2}{3} & \\ \frac{3}{8} & -\frac{7}{3} & \frac{3}{8} \\ & \frac{2}{3} & -\frac{11}{8} \end{pmatrix} & \\
 & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix} & \begin{pmatrix} -\frac{29}{21} & \frac{3}{4} & \\ \frac{8}{21} & -\frac{10}{4} & \frac{8}{21} \\ & \frac{3}{4} & -\frac{29}{21} \end{pmatrix} \\
 & & \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} & \begin{pmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{pmatrix}
\end{pmatrix}.$$

Zur Vollständigkeit notieren wir noch die Dreieckszerlegungen der Matrizen T_j.

$$T_1=\begin{pmatrix} 4 & & \\ -1 & \frac{15}{4} & \\ & -1 & \frac{56}{15} \end{pmatrix}\cdot\begin{pmatrix} 1 & -\frac{1}{4} & \\ & 1 & -\frac{4}{15} \\ & & 1 \end{pmatrix},\qquad T_2=\begin{pmatrix} \frac{11}{3} & & \\ -1 & \frac{71}{22} & \\ & -1 & \frac{715}{213} \end{pmatrix}\cdot\begin{pmatrix} 1 & -\frac{3}{11} & \\ & 1 & -\frac{22}{71} \\ & & 1 \end{pmatrix},$$

$$T_3=\begin{pmatrix} \frac{29}{8} & & \\ -1 & \frac{266}{87} & \\ & -1 & \frac{3509}{1064} \end{pmatrix}\cdot\begin{pmatrix} 1 & -\frac{8}{29} & \\ & 1 & -\frac{87}{266} \\ & & 1 \end{pmatrix},\ T_4=\begin{pmatrix} \frac{76}{21} & & \\ -1 & \frac{113}{38} & \\ & -1 & \frac{7790}{2373} \end{pmatrix}\cdot\begin{pmatrix} 1 & -\frac{21}{76} & \\ & 1 & -\frac{38}{113} \\ & & 1 \end{pmatrix}.$$

(2) $W_j=tridiag(\check{w}_{ij},\bar{w}_{ij},\hat{w}_{ij})$

Analog zum diagonalen Ansatz gelangt man über die Auswertung von $E_j^TT_j=E_j^T(D_j-L_jW_{j-1})$ und der Identifikation

$$\begin{aligned} c_{ij}q_{i-1,j}+m_{ij} &= a_{ij}-l_{ij}\bar{w}_{i,j-1}, \\ m_{i-1,j}q_{i-1,j} &= b_{i-1,j}-l_{i-1,j}\hat{w}_{i-1,j-1}-l_{i+1,j}\check{w}_{i+1,j-1} \end{aligned}$$

auf ein Berechnungsvorschrift für m_{ij},q_{ij} bei Kenntnis von W_{j-1}.
Aber die zweite Bedingung $E_j^TT_jW_j=E_j^TU_j$ liefert jetzt nicht mehr eine Gleichung für ein Diagonalelement von W_j, sondern T_jW_j ist als Produkt von 2 Tridiagonalmatrizen nun pentagonal. Somit setzen sich in $E_j^TT_jW_j$ die Spaltensummen aus maximal 5 Gliedern zusammen, und in einer Gleichung sind die Unbekannten $\hat{w}_{ij},\bar{w}_{i+1,j},\check{w}_{i+2,j}$ gekoppelt.

$$\begin{array}{llllll}
\tilde{q}_{i-1,j}\hat{w}_{ij}+ & & & & & \\
\tilde{a}_{ij}\ \hat{w}_{ij} & + & \tilde{q}_{ij}\bar{w}_{i+1,j}+ & & & \\
c_{i+1,j}\hat{w}_{ij}+ & & \tilde{a}_{i+1,j}\bar{w}_{i+1,j} & + & \tilde{q}_{i+1,j}\check{w}_{i+2,j}+ & \\
 & & c_{i+2,j}\bar{w}_{i+1,j} & + & \tilde{a}_{i+2,j}\check{w}_{i+2,j}+ & \\
 & & & & c_{i+3,j}\check{w}_{i+2,j} & = u_{ij}
\end{array}$$

Will man diesen Weg weiter beschreiten, wären zusätzliche Bedingungen zu definieren und der Aufwand doch erheblich.

6.1.2 NILU mit Zeilensummenkriterium

Zeilensummenübereinstimmung von A und Q heißt

$$\sum_{j=1}^{N} A_{ij} = \sum_{j=1}^{N} Q_{ij} \quad \forall i.$$

Sei wiederrum $E_j^T = (1, 1, ..., 1)$ (gemeinsame Dimension für Blockmatrizen).
Damit berechnet man die Zeilensummen von

$A: \quad L_j E_j + D_j E_j + U_j E_j,$
$Q: \quad L_j E_j + (L_j W_{j-1} + T_j) E_j + T_j W_j E_j.$

Übereinstimmung ist zu erreichen, falls

$$(*) \qquad \begin{aligned} D_j E_j &= (L_j W_{j-1} + T_j) E_j, \\ U_j E_j &= T_j W_j E_j. \end{aligned}$$

Wir betrachten hier nur eine diagonale Struktur, also $W_j = diag(w_{ij})$.

Aus den Bedingungen $(*)$ können folgende Beziehungen abgeleitet werden.

$$\begin{aligned} T_j W_j &= (c_{ij} w_{i-1,j}, \tilde{a}_{ij} w_{ij}, \tilde{q}_{ij} w_{i+1,j}), \\ &= (c_{ij} w_{i-1,j}, (m_{ij} + c_{ij} q_{i-1,j}) w_{ij}, m_{ij} q_{ij} w_{i+1,j}), \\ T_j W_j E_j &= [c_{ij} w_{i-1,j} + (m_{ij} + c_{ij} q_{i-1,j}) w_{ij} + m_{ij} q_{ij} w_{i+1,j}]^T, \\ &\quad n \text{ Zeilensummen bilden Spaltenvektor}, \\ U_j E_j &= [u_{ij}]^T \end{aligned}$$

Damit ist zu jedem j ein tridiagonales Gleichungssystem für w_{ij} zu lösen.

$$c_{ij} w_{i-1,j} + (m_{ij} + c_{ij} q_{i-1,j}) w_{ij} + m_{ij} q_{ij} w_{i+1,j} = u_{ij}, \quad i = 1, 2, ..., n.$$

Die andere Bedingung $D_j E_j = (L_j W_{j-1} + T_j) E_j$ ergibt mit

$$\begin{aligned} D_j E_j &= [c_{ij} + a_{ij} + b_{ij}]^T, \\ L_j W_{j-1} E_j &= [l_{ij} w_{i,j-1}]^T, \\ T_j E_j &= D_j E_j - L_j W_{j-1} E_j = [c_{ij} + a_{ij} + b_{ij} - l_{ij} w_{i,j-1}]^T \end{aligned}$$

die Beziehung

$$c_{ij} + \underbrace{(m_{ij} + c_{ij} q_{i-1,j})} + \overbrace{m_{ij} q_{ij}} = c_{ij} + \underbrace{a_{ij} - l_{ij} w_{i,j-1}} + \overbrace{b_{ij}}.$$

Letztere ist erfüllt bei gleicher Identifikation wie im Spaltensummenkriterium (Vergleich der geklammerten Terme auf beiden Seiten).

Zusammenfassung der 3 Bedingungen:

$$j = 1, 2, ..., m$$

$$\left\{ \begin{array}{rcl} \left.\begin{array}{rcl} m_{ij} & = & a_{ij} - c_{ij} q_{i-1,j} - l_{ij} w_{i,j-1} \\ m_{ij} q_{ij} & = & b_{ij} \\ & & w_{i,0}, q_{0,j} \ \text{ sind Null} \end{array}\right\} & i = 1, 2, ..., n-1, (n), \\ c_{ij} w_{i-1,j} + (m_{ij} + c_{ij} q_{i-1,j}) w_{ij} + m_{ij} q_{ij} w_{i+1,j} = u_{ij}, & i = 1, 2, ..., n, \\ w_{0,j}, w_{n+1,j}, q_{0,j} \ \text{ sind Null.} & \end{array} \right.$$

Ein "Zeilensummenvorkonditionierer" erfordert die Lösung eines tridiagonalen Gleichungssystems, die Elemente der Diagonalmatrix W_j sind also miteinander gekoppelt. Die Berechnungen von m_{ij}, q_{ij} sind in beiden Fällen identisch, also T_j dieselben Matrizen.

Zahlreiche Untersuchungen an Beispielen in [1] führen zur Einschätzung, daß im Rahmen der Vorkonditionierung die Anwendung der tridiagonalen approximativen Inversen zu empfehlen ist sowie auch die NILU-Zerlegung mit Spaltensummen eine robuste Variante darstellt.

Literatur

[1] LEAF, G.K.; MINKOFF, M.; DÍAZ, J.C.: *Nested Block Factorization Preconditioners for Convective-Diffusion Problems in Three Dimensions.* Mathematics for Large Scale Computing (Lecture notes in pure and applied mathematics, vol. 120, ed. J.C.Díaz), New York 1989, pp. 217-263.

[2] HÄMMERLIN, G.; HOFFMANN, K.-H.: *Numerische Mathematik.* Grundwissen Mathematik 7. Springer-Verlag Berlin 1991.

[3] KIELBASINSKI, A.; SCHWETLICK, H.: *Numerische lineare Algebra.* Mathematik für Naturwissenschaft und Technik Band 18, DVW, Berlin 1988.

[4] HACKBUSCH, W.: *Iterative Lösung großer schwach besetzter Gleichungssysteme.* Leitfäden der angewandten Mathematik und Mechanik Band 69. B.G. Teubner Stuttgart 1991.

[5] ZLATEV, Z.: *Computational Methods for General Sparse Matrices.* Math. and Its Appl. Vol.65. Kluwer Academic Publishers London 1991.

[6] GUSTAVSON, F.: *A Survey of Some Sparse Matrix Theory and Techniques.* Jahrbuch Überblicke Mathematik. B.I.-Wissenschaftsverlag Mannheim 1981.

Anschrift:

Dr. Werner Neundorf, stud. cand. Thomas Ortlepp
Technische Universität Ilmenau, Institut für Mathematik
PF 10 0565
D - 98684 Ilmenau
e-mail: neundorf@mathematik.tu-ilmenau.de